Methodological Memo on Measuring Populist Rhetoric on Twitter[1]

In this memo, I present a method for identifying populist rhetoric on Twitter. The method is based on the manual coding technique of holistic grading that was implemented in the Global Populism Database (GPD). Methodologically, this method can be carried out cross-nationally and can be adapted to both populist and non-populist actors.[2] It is also flexible in terms of its coding categories, it minimizes bias through its coding procedures, and it demonstrates both internal reliability and external validity. What sets this method apart from a theoretical standpoint is the holistic approach to the kinds of worldviews (master frames) present in discourse—I incorporate not only populism, but pluralism, technocracy, and neutral rhetoric as well. More precisely, I propose a classification schema that is capable of evaluating how populism appeals to people *relative* to other ways of seeing the political world at two levels of analysis: the master frame and the frame level.

In doing so, this study offers a more precise specification of populist rhetoric in a competitive communication environment. Understanding what kinds of messages resonate in competitive environments is critical in explaining not only vote choice but other kinds of attitudes and behaviors as well.[3] As social media become increasingly important to candidates' campaign strategies, examining how individuals respond to political stimuli on social media offers unparalleled insight into real-world behavior in response to actual campaign messaging. The observational design of this method, while suffering from limitations, also possesses considerable strength by evaluating peoples' actual behavior online in response to real-world campaign messages.

This memo will proceed as follows. First, I describe the strategic advantages of focusing on social media, followed by a review of how this method extends existing scholarship on populism. Then, I describe the different coding categories. Although the main focus of the technique is on rhetoric, I also incorporate relevant dimensions that may be of use to scholars interested in populism beyond or in addition to its rhetorical components, such as the issues political actors mention in their Tweets, the function of the Tweets, and the valence of the Tweets. Next, I describe the coding procedures that I implemented in the original study and which can serve as the basis for scholars interested in performing a similar analysis. I then present preliminary results based on five national-level campaigns where at least one populist candidate ran in 2018 and 2019, including Italy, Mexico, Brazil, Colombia, and Spain (Cassell n.d.) to show how the

---

[1] Author: Kaitlen Cassell, PhD Candidate at Vanderbilt University. Kaitlen.j.cassell@vanderbilt.edu. This memo represents a short methodological summary of the method used to measure populism in Twitter that is the focus of my dissertation.

[2] But see Aslanidis (2018) for a novel method of text analysis called semantic text analysis that utilizes subject-verb-object (SVO) clauses as the unit of analysis. Aslanidis's work offers an alternative way to measure populism, and is translatable down to the grassroot level (not just among political actors). For the purposes of this study, Aslanidis's method did not fit with the objective of my study, which was to apply different master frames and disaggregated frames which required interpreting the Tweet as a whole rather than individual clauses.

[3] Many scholars focus primarily on the outcome of populist vote choice (Aguilar and Carlin 2018; Akkerman, Zaslove, and Spruyt 2017; Bakker, Rooduijn, and Schumacher 2016; Van Hauwaert and Van Kessel 2017; Hawkins, Rovira Kaltwasser, and Andreadis 2016). Though vote choice is a critical outcome, populism's success at the polls is only one metric of the public's receptivity to the populist agenda, and a noisy one at that—vote choice is determined by any number of factors. As a result, vote choice in isolation likely leads to an under-reporting of populist identification, whereas engagement focuses more directly on how receptive the public is to the populist message.

method works in practice.[4] I conclude with a brief discussion of the relative merits of the method and its potential for future research.

Why Social Media?

Social media have become increasingly important as more communication and participation move from traditional political spaces to online ones (Gil de Zúñiga 2012; Zamora Medina and Zurutuza Muñoz 2014). Despite its growing importance, social media is comparatively understudied in the populist literature. Yet, it offers several strategic advantages to politicians as well as methodological advantages to scholars.

Of the many ways that social media can benefit a political actor, two are particularly relevant. First, social media affect the way that candidates communicate with their followers. For one, social media offer political actors the opportunity to spread their messages to a broad audience. Furthermore, social media represent unmediated, real-world conversations between political actors and their followers, which may disproportionately benefit populist political actors (Barr 2009; Enli 2017; Enli and Rosenberg 2018). Additionally, social media are highly dynamic: political actors not only post messages, but announce events, videos of speeches/rallies, and spread campaign ads, among other functions. Second, social media can alter outcomes that political actors are interested in achieving. Some populists have gone as far as to attribute part of their success to social media—in his victory speech on July 1, 2018, the current president of Mexico, Andrés Manuel López Obrador (AMLO), expressed his gratitude "to the blessed social media."[5] Scholars have also provided evidence that candidates set the media's agenda with their posts (Enli 2017; Graham et al. 2014).

From a measurement perspective, social media offer a way to measure engagement that is based on actual political behavior as a response to real candidate rhetoric rather than artificial responses to artificial stimuli. From a methodological standpoint, social media posts have the potential to generate a large N both across cases and over time. Relatedly, the availability of social media posts allows scholars to implement random sampling techniques, decreasing the possibility of selection bias. I evaluate Tweets over other social media due to the standardized maximum length of posts. In texts of 280 characters or less, Tweets represent fragments of discourse that are easier to code compared to texts of unlimited length (Engesser et al. 2017).

Existing Research

There are three relevant strands of populist literature that inform the technique: 1) identifying populists/non-populists based on their communication, 2) studies that classify the particular populist frames that political actors use, and 3) studies that focus on which frames generate engagement.[6] The first of these literatures, *identifying populists based on their rhetoric,* forms

---

[4] For each country, I evaluate a random sample of the Tweets of all candidates that pass a 10% vote share threshold for a total of eighteen candidates: eight who use populist rhetoric consistently in their Tweets (i.e., populists) and ten who do not (i.e., non-populists) for an N of 1,060 Tweets.

[5] https://twitter.com/diazbriseno/status/1013640626703945728

[6] For the purposes of this memo, I focus exclusively on online engagement. However, scholars have done several experimental studies that test how respondents engage with populist frames via other forms of participation or behaviors (see, e.g., Busby, Hawkins, and Gubler 2018; Hameleers, Bos, and de Vreese 2018; and Hameleers and Schmuck 2017).

the foundation of subsequent studies. More concretely, scholars identify political actors who use populist rhetoric (referred to henceforth as populists) and candidates that do not (non-populists), generally assigning these actors a score that represents how populist they are (Hawkins and Castanho Silva 2018; Rooduijn and Pauwels 2011). Scholars have built on this work to identify different kinds of populist frames in politicians' communication and the frequency with which these frames are used (Casero-Ripollés, Sintes-Olivella, and Franch 2017; Cranmer 2011; Ernst et al. 2017). This *classification literature* complements the above studies by providing a substantive look at the rhetorical components making up an actor's populist "score." More recently, a few scholars have begun testing the association between different kinds of frames and measurable aspects of online *engagement,* such as the number of likes or re-Tweet a message receives.

The technique described here offers several advantages over previous ones for measuring populism on Twitter. Previous work *identifying populist rhetoric* typically focuses on traditional forms of communication like speeches (Hawkins and Castanho Silva 2018), manifestos (March 2018; Rooduijn and Pauwels 2011), or interviews (Grbesa and Salaj 2018). The main advantage of the method I propose is that it extends populism identification studies to social media, the benefits of which I describe above. It also provides additional information beyond a populists' score, such as what frames they use and how often.

The primary advantage of this technique over other *classification studies* is that it incorporates a more complete picture of rhetoric. Scholars often focus exclusively on populist rhetoric in isolation. For example, Casero-Ripollés, Sintes-Olivella, and Franch (2017) find that 51.9% of Podemos's tweets use populist frames—but what about the other 48.1%? If populist candidates use only a slim majority of populist frames, what other frames are they using, and what frames are non-populist candidates using? The technique I propose can answer these questions, allowing scholars to evaluate 100% of a candidate's messages instead of only the populist ones. I consider the extended rhetorical classification schema as the primary theoretical contribution of this method.

In addition to extending the scope of rhetoric, I also focus on a narrower, theory-driven set of populist frames than some scholars consider.[7] My classification schema uses three populist frames: "pro-people," "anti-elite," and "dispositional blame attribution."[8] These three elements correspond to the ideational theory's three necessary and sufficient elements of populism (Hawkins and Rovira Kaltwasser 2018). Existing studies focus on anywhere from four to six frames.[9] Earlier iterations of this technique demonstrated that other frames either were biased

[7] I derive the non-populist frames (technocratic, pluralist, and neutral, specifically) from a combination of existing studies (particularly for the populist frames), codebooks (Hawkins 2019), theoretical studies, and survey measures (particularly for the pluralist and technocratic frames, for which there is less literature). Once I created an initial schema, I applied it to a random sample of 120 Tweets, resulting in a final schema of thirteen unique frames.

[8] I refer to Manichean discourse as dispositional blame attribution, but the underlying sentiment behind these two concepts, as they are used here, is substantively the same.

[9] Casero-Ripollés et al. (2017), Cranmer (2011), and Engesser et al. (2017) incorporate the theme of "the people" and "the elite" mirroring two of the three defining components of populism, per the ideational theory (Hawkins et al. 2018). However, Casero-Ripollés et al. proposes a single frame for "the people," Engesser et al. two frames ("sovereignty of the people" and "advocacy for the people"), and Cranmer three frames ("advocacy for," "accountability to," and "the legitimacy of the people"). All three articles include a frame for "attacking the elite" and "exclusion of outgroups." At this point, the framing schemas diverge more substantially. Engesser et al. proposes a frame that invokes "the heartland;" Casero-Ripollés et al. a frame for "narrative of a crisis," and Cranmer a frame for "homogeneity or threat."

towards right-wing populism, occurred extremely infrequently (less than 1% of the sample), or were not distinct from the three final frames used in the study.

Though this method can be used to classify populist frames for descriptive purposes, it can also be used to quantify the effect of particular attributes of a message (described below) on a relatively new dependent variable: online engagement. This method allows scholars to leverage the number of likes and re-Tweets as quantifiable dependent variables, permitting various estimation techniques.[10] Very few studies have made use of this promising outcome. Bobba (2019) and Bobba and Roncarolo (2018) are the only scholars I am aware of to test the effects of message characteristics on the likeability of Tweets. While these studies contribute to the nascent literature on the engagement potential of populist messages, their generalizability is limited to one election in one country and only populist frames.

In contrast, the technique I propose improves upon existing studies in a few key ways. First, this method incorporates both likes and re-Tweets for a more comprehensive picture of engagement. As described above, the study also focuses on a broader spectrum of rhetoric: doing so allows for a richer comparison of the rhetorical strategies of *non-populists*, who are often evaluated only by the extent to which they use populist rhetoric. Finally, the method can be applied across countries and candidates, allowing for greater generalizability within and between cases.

A better way to think about this technique is that it combines these related literatures while simultaneously working to advance them. The technique I propose can be used to classify how populist a candidate is based on the percentage of populist frames they use on a form of communication that has become increasingly important in campaigns specifically and political communication more generally: social media. Most obviously, it allows for classification of the different kinds of frames political actors use and at what frequency, including but not limited to populism. Finally, it allows scholars to make inferences about which of these frames are associated with higher or lower levels of online engagement from a cross-national perspective.

Description of the Method

Concepts & Measures

One of the chief advantages of this method is its flexibility. In other words, it can be used to classify actors as populist or not populist, to classify a broad spectrum of rhetoric, and to measure real-world political behavior. I incorporate the following coding dimensions, which can be tailored to a wide variety of research questions. Additional details on each category, including examples, are available in the Appendix.

- ▪ **Actors**: *Who does the tweet reference (implicitly or explicitly)?* Who is the one that is doing the action and who is the one receiving it

Actor
"the people" *(populist and non-populist interpretations)*

---

[10] Depending on the web-scraping method implemented in collecting Tweets. I used a combination of Twlets, a paid service for downloading recent Tweets (the last 3,200), and a chrome-based tool (DataMiner) with pre-developed packages to get around Twitter's API, which only allows you to scrape the last 3,200 Tweets. Note that when a speaker re-Tweeted someone else's message, it was excluded from the analysis since it did not represent the speaker's words.

"the elite" *(populist and non-populist interpretations)*	
In-group or out-group	
The candidate, their party, or members of their party	
The opposition	
The media	
No actors	

- **Issue**: *What is the main topic of the tweet?* I employ broad 12 issue categories, adapted and simplified from Casero-Ripollés, Sintes-Olivella, and Franch (2017).

Issue	Description
Economy	Tweets including subjects such as jobs, unemployment, salaries, deficit, public spending, debt, crisis, taxes, entrepreneurship, contracts, self-employed people, agricultural policy, and so on. This is a somewhat narrow category that should refer explicitly to the economic realm.
Social policy	Tweets including subjects such as pensions, health, education, the welfare state, poverty, social justice, equality/inequality (including gender-based violence), housing, immigration, childbirth, drug rehabilitation, and so on. This is a broader category that encompasses some economic-adjacent issues (inequality, welfare) that affect people.
Culture, media, and sport	Tweets including subjects related to cultural industries (cinema, literature, art, mainstream media, social media, etc.) and sport.
Science, technology, the environment, and infrastructure	Tweets including subjects related to research and development, network infrastructure (such as fiber optic, ADSL, or Wi-Fi), transportation infrastructure (railway, airports, roads, etc.), pollution, flora and fauna protection, climate change, and so forth.
Terrorism, crime, and insecurity	Tweets related to terrorism in all its forms and crime/criminal activity or general concerns about insecurity.
Foreign affairs	Tweets alluding to the European Union, the United States, international relations, or other parts of the world.
Corruption and democratic regeneration	Tweets including subjects concerning political corruption and/or democratic aspects that need to be renewed or removed, like changes in electoral law, putting an end to the establishment and the privileges of the political class, and so on.
Political strategy in office	Tweets including subjects concerning the intention of the **candidate if they were to win office** (i.e., not specific to the campaign period itself). For example, forming a certain type of government or possible (or impossible) government pacts/coalitions in the future. Additionally, if the candidate Tweets about multiple issue positions (the economy and social positions), classify it as political strategy.
Campaign organization and strategy	Tweets including subjects concerning the candidate **during the campaign period**. This can include questionnaires, surveys, information, analysis, and assessment of electoral results, or Tweets referring to the action of voting. It can also refer to Tweets about the running of the campaign and the organization of events, like rallies, meetings, political events, and media appearances by the candidates (more specific), or Tweets exalting the importance of party unity and exhorting sympathizers to join the party and earn victory (more broadly).
Immigration	Tweets about the topic of immigration
Regional politics	Tweets relating to political subdivisions such as particular regions, states, etc. Note: this should not be used whenever a candidate talks about a particular city; it is more about the distribution of power within a country, such as the secession movement in Catalonia, Spain, or urban vs. rural politics.
No subject or Other	Tweets that do not have a defined subject or that include expressions of courtesy (acknowledgments, etc.) or Tweets referring to the personal life of political agents. Tweets that cannot be placed in the above categories.

▪ **Function:** *What is the speaker trying to accomplish?* This study utilizes 8 function categories, also adapted and simplified from Casero-Ripollés, Sintes-Olivella, and Franch (2017).

Function	Description
Agenda and organization of political actions (including media appearances)	Tweets containing information on specific campaign actions in which the time and place are specified. **This should take place either in the near future, or be in progress at the time the Tweet is sent**. Tweets sharing links to a journalistic interview or TV show.
Electoral program	Tweets on future political proposals or program proposals. This should be somewhat specific—not just vague intonations of making the country better.
Management of political achievements	Tweets extolling or praising the achievements of the party and/ or leader. This could also include things like endorsements or responses to polls/early election predictions.
Criticizing opponents	Tweets containing direct or indirect attacks on other candidates, political parties, other leaders (past or present) or other ideologies more broadly.
Participation and mobilization	Tweets aimed directly at increasing support/votes **during the campaign**. This can include the mention of general campaign events (we were in XX city this morning), but the reference should be somewhat vague. Followers would not know where to go or what type of event based just on this Tweet alone (in contrast to function 1).
Personal life/ backstage or Manners/Protocol	Tweets where particularly the leaders show or talk about things from their private lives (leisure, hobbies, sport, etc.) or from backstage at political events or from the campaign. Tweets of thanks, sympathy, greetings, special occasions, and so on.
Entertainment or Humor	Tweets encouraging community building around the party or the leader with an entertainment-based focus, or Tweets containing memes, jokes, or other humorous resources.
Others	Tweets that cannot be placed in the above categories.

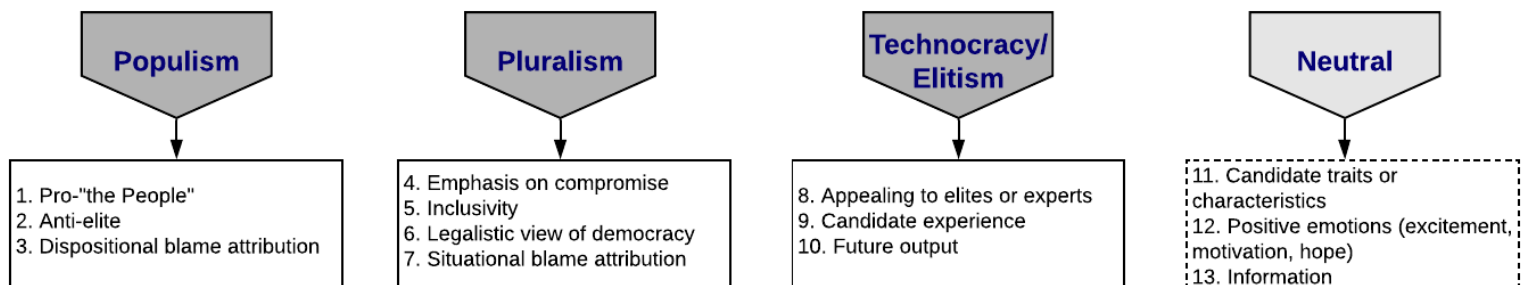▪ **Valence**: *What kind of language does the Tweet use?*

Valence
Predominantly **positive** language
Neutral language, or equally positive and negative language
Predominantly **negative** language

▪ **Master frame**: *How do people see the world in terms of who should hold power?* The 4 master frames analyzed in Cassell (n.d.) include populism, pluralism, technocracy, and neutral messages. Master frames can be differentiated from one another and measured in discourse according to 1) who should be in charge and 2) where their legitimacy comes from (Caramani 2017; Hawkins 2010).

Master Frame	Who is in Charge	Source of Legitimacy
Populism	Delegated leaders who represent the "will of the people; often portrayed as a "common man" or one of "the people"	Understanding and embodying the putative will of "the people"
Pluralism	Political leaders within formal party structures who represent diverse interests	Formal democratic representation (winning a fair election)
Technocracy	Rational experts; individuals who have a particular claim to authority based on membership into a specific group of authorities	Rational speculation; being distant/distinct from "the people"

Neutral	This is a "catch-all" category for Tweets that cannot in and of themselves be classified into just one of the above master frames. This is usually due to ambiguity—neutral frames can apply to a number of different world views, and this ambiguity necessitates its own category

- **Disaggregated frame**: *What are the individual rhetorical elements that uniquely make up a master frame?* Using populism as an example, the ideational theory of populism holds that three components are necessary and sufficient to define populism: pro-people, anti-elite, and a Manichean (dualistic) world view (Hawkins et al. 2018). At this lower level of rhetoric, Tweets were classified as one of these elements. In total, I devise a schema with 13 disaggregated frames: 3 populist, 3 pluralist, 4 technocratic, and 3 neutral. Details on each frame are presented in the Appendix.

Populism

1. Pro-"the People"
2. Anti-elite
3. Dispositional blame attribution

Pluralism

4. Emphasis on compromise
5. Inclusivity
6. Legalistic view of democracy
7. Situational blame attribution

Technocracy/ Elitism

8. Appealing to elites or experts
9. Candidate experience
10. Future output

Neutral

11. Candidate traits or characteristics
12. Positive emotions (excitement, motivation, hope)
13. Information

Other categories include the perceived strength of the frame, the difficulty in classifying the frame, and for frame 10) output, the specificity of the issue (from vague to the identification of a concrete issue and projected outcome). Scholars could also add their own categories to suit their research needs.

Coding Procedures

Four coders, three undergraduates, and the author, coded the Tweets. Undergraduate coders were trained with a codebook and two practice samples. To decrease bias and enhance the validity of the coding process, coders were not told the purpose of the study (i.e., to examine populist, pluralist, and technocratic rhetoric), or what concepts the master frames corresponded to. The master frames were described in-depth but labeled as "Type A," "Type B," etc. While it is likely that as political science students, the coders recognized some of the conceptual underpinnings of the master frames, this procedure was put in place to ensure that no explanation was privileged. Had I told the coders this was a study about populism, they may have been intentionally or unintentionally biased towards finding populist frames. Intercoder reliability tests show that, if anything, coders were biased against populism (the final coding contained more populist frames than the undergraduate coders identified).

Each Tweet in the final sample was classified according to a three-stage procedure. First, the Tweets were **de-identified (stage 1)** and subsequently translated,[11] which included masking

---

[11] Tweets were presented both in their original language, as well as an English translation using Google Translate. One undergraduate coder fluently spoke all three languages in this study (Spanish, Italian, and Portuguese), and assisted in correcting the Google Translated output when necessary. The author also made similar corrections for Spanish-speaking cases.

the candidate's identity, party/political affiliation, and date of the Tweet.[12] This stage also included adding contextual clues for the coders that did not reveal speaker identities but provided enough information for coding.

Example 1:
- Haddad, Lula's candidate, visits the Technological Park of Sao Jose dos Campos. #HaddadPresidente https://t.co/n6oSUgb9WA
- *[party candidate], [party leader]*'s candidate, visits the Technological Park of São Jose dos Campos. *#[party candidate]*forPresident *[link removed]*

Example 2:
- THE TG1 STRIKES AGAIN! We ask Tg1 not to participate in Renziana propaganda. In today's report by Maesano, the lies of Renzi and the parliamentarian Alessia Rotta are reported without contradiction about the decrease of the tax burden. https://goo.gl/tf3m2b pic.twitter.com/nVfzSTXr2Q
- THE TG1 STRIKES AGAIN! We ask Tg1 not to participate in *[opposition party]* propaganda. In today's report by Maesano, the lies of *[leader of the opposition]* and the *[parliamentary member of opposition]* are reported without contradiction about the decrease of the tax burden. *[link removed] [photo removed]*

De-identification was carried out by the author or one trained multi-lingual undergraduate RA. This division ensured that, in Stage 2, one undergraduate coder was blind to the speaker's identity and information, while the other coder was not. All Tweets were then randomized (across candidates but within countries for language purposes) and translated. In some cases, the coders needed to view the media attached to the Tweet in order to accurately code it, thus exposing the candidate's identity. Media that met this standard include threads or consecutive Tweets (Graham et al. 2014; Welp and Ruth 2017), short videos, news articles, links to longer posts, and infographics. The inclusion of non-text in the coding decision is an important divergence from some studies (see, e.g., Bobba and Roncarolo 2018). The decision to include these items was made on the basis that important contextual clues are regularly found in non-text. Media that did not meet this standard (and were removed) were those that did not contain useful context, including photos or videos of the speaker or the crowd at campaign events, images that duplicated the text of the Tweet, or links to videos that were more than a couple of minutes long. In the latter instance, the de-identifiers in Stage 1 provided a short description of the media without indicating the speaker's identity.

Next, **independent classification by two coders (stage 2)** took place. Coders were unable to view each other's answers. Coders provided detailed explanations (1-4 sentences) for each coded Tweet to justify their chosen classification, which was factored into the final coding decision. Lastly, the Tweets were subject to **final review by the author (stage 3)** to reconcile potential divergences. The author reviewed the two coding decisions and relied on a combination of the coders' explanations and meetings with the coders to make the final coding determination. In the event of an incorrect coding determination (such as if the coder misunderstood the text),

---

[12] De-identification, randomization, and translation took place in Excel. Subsequent stages (2 and 3) were conducted in Vanderbilt's data management software, Redcap, which minimized user error by using drop-down selection menus instead of manual data entry.

the author provided ongoing comments explaining the coding decision, but coders did not change their coding decisions once the initial determination had been made so as not to bias the results.

The biggest downside of this method is that the coding process is time-intensive. Training coders requires extensive feedback after each training set, but this feedback decreases over time as the coders become more experienced. De-identification can also be cumbersome, though this stage would be expedited if the coder is familiar with the main political actors in the country. It is also worth noting that de-identification is not absolutely necessary. I implemented this procedure to attempt to limit coders' biases, but most existing studies do not engage in similar practices, particularly those that exclusively aim to identify populist rhetoric. The advantages of the method, particularly its reliability, outweigh the time costs, but should nonetheless be taken into consideration.

Preliminary Results

I apply the classification schema to the campaign messages of five national-level campaigns where at least one populist candidate ran in 2018 and 2019, including Italy, Mexico, Brazil, Colombia, and Spain (Cassell n.d.). For each country, I evaluate a random sample of the Tweets for all candidates that pass a 10% vote share threshold, totaling eighteen candidates (N=1,060). I present the results for master frames and disaggregated frames below. The below figures rank each candidate in terms of the percentage of Tweets that are classified as populist, while also indicating the percentages of pluralist, technocratic, and neutral frames. Blue text indicates candidates/parties who use >20% of populist frames.

Master Frames
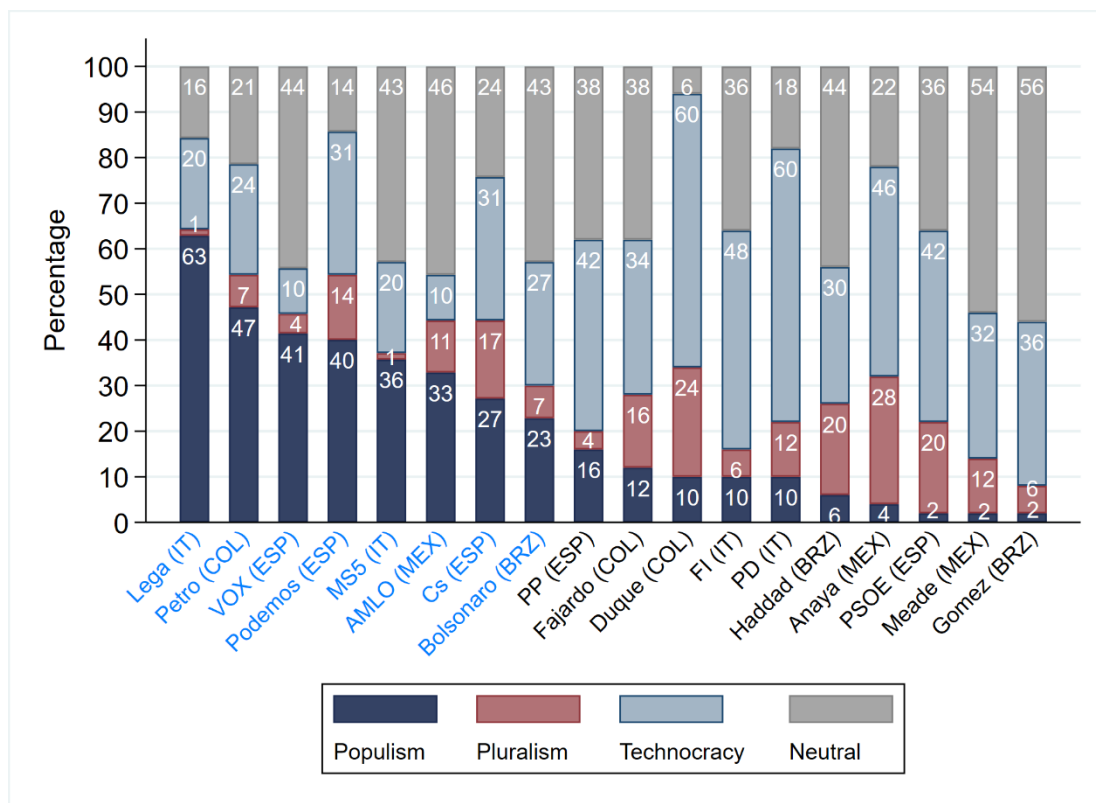
Disaggregated Frames
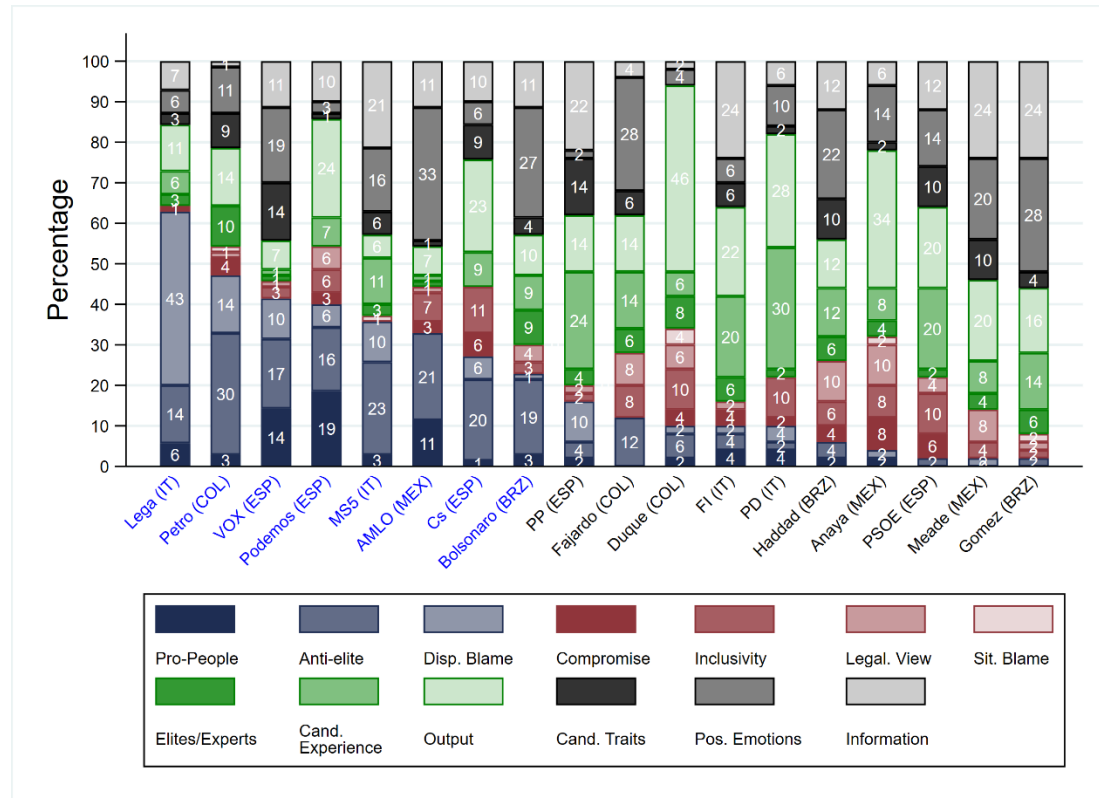
These figures dispel the notion that populist candidates exclusively or even primarily employ populist rhetoric, highlighting the need to focus on the other rhetoric they use. Besides Lega, no other candidate/party uses a majority of populist frames—even candidates that are generally seen as quintessential populists such as Podemos, MS5, and AMLO use only 33 to 40% populist frames. With few exceptions, populists on Twitter are using less pluralistic and technocratic messages than their non-populist counterparts (*p<.01* each), although populists and non-populists are not statistically differentiable in their use of neutral frames. More generally, the data reflect a wide range of frames used, both across and within candidates. Candidates do not just stick to one worldview or even one frame within a worldview—no candidate uses even a majority of a single frame.

Method Validity

This method displays high levels of internal reliability and external validity. The intercoder reliability for the entire sample, presented using Krippendorff's alpha, is .66 at the master frame level and .63 at the disaggregated frame level. Acceptable levels of reliability generally range between .6 and .8 (Krippendorff 2018). In terms of external validity, this method performs well compared to existing measures of populism across types of communication. To evaluate the method of discourse analysis that I designed and used in this project, I compare it to other existing databases for triangulation purpose. I compare three existing measures of

populism: The Global Populism Database, the Chapel Hill Expert Survey, and the Negative Campaigning Comparative Expert Survey.

The Global Populism Database classifies how populist a candidate is, ranking candidate and politicians' scores along a 0-2 scale with four classification benchmarks: not populist (0-0.49); somewhat populist (0.5-0.99); populist (1-1.49); very populist (1.5-2).[13] Note that the GPD classifies political candidates/actors only, not parties.

The Chapel Hill Expert Survey (CHES) relies on the opinion of political experts. In 2017, CHES asked experts to classify parties according to two dimensions: the people versus the elite and the salience of anti-elite rhetoric, each on a 0-10 scale with 0 indicating a non-populist perception of this party, and 10 indicating a populist response.[14]

The Negative Campaigning Comparative Expert Survey (NEGex) also relies on political experts. NEGex asks experts to rate candidates on three populist dimensions: identification with the people, respect for opponents (which I refer to as anti-elite), and simplicity of the message.[15] Note that NEGex classifies political candidates/actors only, not parties. I impose the following cut-off points that are based on the answer choices of experts: not populist (all scores are below 3.0), somewhat populist (only one element of populism exceeds 3.0), and populist (at least two elements of populism exceed 3.0).

For ease of interpretation, I use the following classification benchmarks in my data set: not populist (0-2.49); somewhat populist (2.5-4.99); populist (5-7.49); and very populist (7.5-10). Not all parties/candidates are present in each data set.

The below table indicates that, using simple populist-or-not dichotomous terms, the Twitter Rhetoric Database is consistent with other measures in **29/34** observation points (where some candidates/parties were evaluated by more than one data set and thus had multiple comparison points). Every political actor in the analysis was evaluated by at least one of these comparative data sets. Discrepancies are noted in red text. The mismatch with Anaya (MEX) compared to the GPD is likely a product of the specific sample used—one of the three speeches that determined his score is driving up the results with a populist score of 1.4; the other two speeches were coded as 0 and .55, respectively. Notably, NEGex finds Anaya to be on the low end of the index, which is consistent with my findings.

FI is somewhat more puzzling: in this sample, FI uses only 8% populist frames. This result is consistent with Bobba and Roncarolo (2018), who classify only 8.1% of FI's Tweets as populist. The divergence could be a product of the enigmatic figure of Berlusconi, who may appear populist without using a significant amount of populist frames. It is also worth noting that Berlusconi just makes the "somewhat populist" benchmark of The Global Populism database.

---

[13] More information can be found at https://www.theguardian.com/world/2019/mar/06/how-we-combed-leaders-speeches-to-gauge-populist-rise; the data can be found at https://populism.byu.edu/Pages/Data.

[14] The people versus elites question asks: "Some political parties take the position that "the people" should have the final say on the most important issues, for example, by voting directly in referendums. At the opposite pole are political parties that believe that elected representatives should make the most important political decisions. Where do the parties fall on this dimension?" The anti-elite rhetoric question asks: "Next, we would like you to think about the salience of anti-establishment and anti-elite rhetoric for a party. How important was the anti-establishment and anti-elite rhetoric to the parties in their public stance?" Chapel Hill Expert Survey 2017 Codebook: www.chesdata.eu.

[15] The surveys ask experts: *And how would you say that the following statements apply to {candidate}? In your opinion, {candidate} might be someone who...1)* Identifies with common people, 2) Uses informal style, popular language, and 3) Uses anti-establishment/elite rhetoric. The answer choices are 0-4, from strongly disagree to strongly agree. The scores presented are averages from all experts that evaluated a particular actor. Data and documentation can be found at https://www.alessandro-nai.com/negex-data.

Leader:	AMLO (MEX)	Anaya (MEX)	Meade (MEX)	Duque (COL)	Petro (COL)	Fajardo (COL)	Bolsonaro (BRZ)	Haddad (BRZ)	Gomez (BRZ)
% of Populist Frames Used (This study)	Populist (33%)	Not populist (4%)	Not populist (2%)	Not populist (8%)	Populist (46%)	Not populist (12%)	Populist (21%)	Not populist (6%)	Not Populist (2%)
Speech Results (Hawkins)	Very populist (1.6)	Somewhat populist (.60)	Not populist (.01)	Not populist (.075)	Somewhat populist (.95)	Not populist (.0375)	Somewhat populist (.5)	*Not rated*	*Not rated*
NEGex (Nai)— pro-people, anti-elite, simple messaging	Populist (3.91, 3.73, 3.82)	Not populist (0.83, 1.17, 1.25)	*Not rated*	Not populist (1.33, 2.57, 1.0)	Populist (3.88, 3.88, 3.88)	*Not rated*	Populists (2.33, 3.13, 3.33)	Not populist (2.75, 2.13, 2.63)	Not populist (2.6, 2.8, 2.6)

Leader:	FI (IT)	M5S (IT)	PD (IT)	LN (IT)	PSOE (ESP)	Vox (ESP)	PP (ESP)	Podemos (ESP)	C's (ESP)
% of Populist Frames Used (This study)	Not populist (8%)	Populist (36%)	Not populist (6%)	Populist (63%)	Not populist (2%)	Populist (41%)	Not populist (12%)	Populist (40%)	Populist (26%)
Speech Results (Hawkins)	Somewhat populist (.5) **Berlusconi	Populist (1.15) **Di Maio	Not populist (.1) ** Renzi	Populist (1.05) ** Salvini	*Not rated*	*Not rated*	*Not rated*	*Not rated*	*Not rated*
NEGex (Nai)—anti-elite, pro-people, simple messaging	Not populist (2.22, 2.89, 1.44) **Berlusconi	Somewhat populist (2.60, 2.80, 3.0) **Di Maio	Somewhat populist (1.71, 3.57 1.71,) ** Renzi	Populist (3.38, 3.5, 3.38) ** Salvini	Not populist	Populist (2.79, 3.29, 3.29) **Santiago Abascal	Not Populist (1.36, 2.21, 0.4) **Pablo Casado	*Not rated*	*Not rated*
CHES (people vs. elites; anti-elite salience)	Somewhat populist (3.75; 3.61)	Very populist (9.75; 10)	Not populist (2.75; 2.46)	Very populist (7.83; 7.85)	Not populist/ somewhat populist (3.5; 2.1)		Not populist (.78; 1)	Very populist (8.78; 8.64)	Somewhat populist (3.65; 5.38)

Conclusion

This method of textual analysis is flexible, mitigates bias, and has high internal reliability and external validity. The identification of populist rhetoric on Twitter is ripe for further application given the integral role social media plays in today's political climate and could be applied to bigger datasets, longer time periods, and even automation. The technique could also be adapted to other social media posts such as Facebook or even Whatsapp.

References

Aguilar, Rosario, and Ryan E. Carlin. 2018. "Populist Voters: The Role of Voter Authoritarianism and Ideology." In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, eds. Kirk Hawkins, Ryan Carlin, Levente Littvay, and Cristobal Rovira Kaltwasser. New York: Routledge, 396–418.

Akkerman, Agnes, Andrej Zaslove, and Bram Spruyt. 2017. "'We the People' or 'We the Peoples'? A Comparison of Support for the Populist Radical Right and Populist Radical Left in the Netherlands." *Swiss Political Science Review* 23(4): 377–403.

Aslanidis, Paris. 2018. "Measuring Populist Discourse with Semantic Text Analysis: An Application on Grassroots Populist Mobilization." *Quality and Quantity* 52(3): 1241–63.

Bakker, Bert N., Matthijs Rooduijn, and Gijs Schumacher. 2016. "The Psychological Roots of Populist Voting: Evidence from the United States, the Netherlands and Germany." *European Journal of Political Research* 55(2): 302–20.

Barr, Robert R. 2009. "Populists, Outsiders and Anti-Establishment Politics." *Party Politics* 15: 29–48.

Bobba, Giuliano. 2019. "Social Media Populism: Features and 'Likeability' of Lega Nord Communication on Facebook." *European Political Science* 18(1): 11–23. https://doi.org/10.1057/s41304-017-0141-8.

Bobba, Giuliano, and Franca Roncarolo. 2018. "The Likeability of Populism on Social Media in the 2018 Italian General Election." *Italian Political Science* 13(1): 51–62.

Busby, Ethan C, Kirk Hawkins, and Joshua R Gubler. 2018. *Speaking Populism: A Framing Experiment to Test the Activation of Populist Attitudes*.

Casero-Ripollés, Andreu, Marcal Sintes-Olivella, and Pere Franch. 2017. "The Populist Political Communication Style in Action: Podemos's Issues and Functions on Twitter During the 2016 Spanish General Election." *American Behavioral Scientist* 61(9): 986–1001.

Cassell, Kaitlen J. 2019. *The Power of Populism: The Effectiveness of Populist Rhetoric in Generating Online Engagement*.

Castanho Silva, Bruno. 2018. "Populist Success: A Qualitative Comparative Analysis." In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, eds. Kirk Hawkins, Ryan Carlin, Levente Littvay, and Cristobal Rovira Kaltwasser. New York, 279–93.

Cranmer, Mirjam. 2011. "Populist Communication and Publicity: An Empirical Study of Contextual Differences in Switzerland." *Swiss Political Science Review* 17(3): 286–307.

Engesser, Sven, Nicole Ernst, Frank Esser, and Florin Büchel. 2017. "Populism and Social Media: How Politicians Spread a Fragmented Ideology." *Information, Communication & Society* 20(8): 1109–26. https://www.tandfonline.com/doi/full/10.1080/1369118X.2016.1207697.

Enli, Gunn. 2017. "Twitter as Arena for the Authentic Outsider: Exploring the Social Media Campaigns of Trump and Clinton in the 2016 US Presidential Election." *European Journal of Communication* 32(1): 50–61.

Enli, Gunn, and Linda Therese Rosenberg. 2018. "Trust in the Age of Social Media: Populist Politicians Seem More Authentic." *Social Media and Society* 4(1).

Ernst, Nicole et al. 2017. "Extreme Parties and Populism: An Analysis of Facebook and Twitter across Six Countries." *Information Communication and Society* 20(9): 1347–64.

Gil de Zúñiga, Homero. 2012. "Social Media Use for News and Individuals' Social Capital,

Civic Engagement and Political Participation." *Journal of Computer-Mediated Communication* 17(3): 319–36.

Graham, Todd, Dan Jackson, and Marcel Broersma. 2014. "New Platform, Old Habits? Candidates' Use of Twitter during the 2010 British and Dutch General Election Campaigns." *New Media and Society* 18(5): 765–83.

Grbesa, Marijana, and Berto Salaj. 2018. "Textual Analysis: An Inclusive Approach in Croatia." In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, eds. Kirk Hawkins, Ryan Carlin, Levente Littvay, and Cristobal Rovira Kaltwasser. New York: Routledge, 67–89.

Hameleers, Michael, Linda Bos, and Claes de Vreese. 2018. "Framing Blame: Toward a Better Understanding of the Effects of Populist Communication on Populist Party Preferences." *Journal of Elections, Public Opinion and Parties* 28(3): 380–98.

Hameleers, Michael, and Desirée Schmuck. 2017. "It's Us against Them: A Comparative Experiment on the Effects of Populist Messages Communicated via Social Media." *Information, Communication & Society* 20(9): 1425–44. https://doi.org/10.1080/1369118X.2017.1328523.

Van Hauwaert, Steven, and Stijn Van Kessel. 2017. "Beyond Protest and Discontent : A Cross-National Analysis of the Effect of Populist Attitudes and Issue Positions on Populist Party Support." *European Journal of Political Research*: 1–25.

Hawkins, Kirk, Ryan Carlin, Levi Littvay, and Cristóbal Rovira Kaltwasser. 2018. *The Ideational Approach to Populism: Concept, Theory, and Method*. Routledge.

Hawkins, Kirk, and Bruno Castanho Silva. 2018. "Textual Analysis: Big Data Approaches." In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, eds. Kirk Hawkins, Ryan Carlin, Levente Littvay, and Cristobal Rovira Kaltwasser. New York: Routledge, 27–48.

Hawkins, Kirk, and Cristóbal Rovira Kaltwasser. 2018. "Introduction: The Ideational Approach." In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, eds. Kirk Hawkins, Ryan Carlin, Levente Littvay, and Cristobal Rovira Kaltwasser. New York: Routledge, 1–24.

Hawkins, Kirk, Cristóbal Rovira Kaltwasser, and Ioannis Andreadis. 2016. *The Activation of Populist Attitudes: Evidence from Contemporary Chile and Greece*.

Krippendorff, Klaus. 2018. *Content Analysis: An Introduction to Its Methodology*.

March, Luke. 2018. "Textual Analysis: The UK Party System." In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, eds. Kirk Hawkins, Ryan Carlin, Levente Littvay, and Cristobal Rovira Kaltwasser. New York: Routledge, 49–66.

Rooduijn, Matthijs, and Teun Pauwels. 2011. "Measuring Populism: Comparing Two Methods of Content Analysis." *West European Politics* 34(6): 1272–83.

Welp, Ranina, and Saskia Ruth. 2017. "Presidentas Twitteras: The Social Media Use of Cristina Fernández de Kirchner and Dilma Rousseff." In *Women, Politics, and Democracy in Latin America*, ed. T. Došek et Al. , 131–49.

Zamora Medina, Rocío, and Cristina Zurutuza Muñoz. 2014. "Campaigning on Twitter: Towards the 'Personal Style' Campaign to Activate the Political Engagement During the 2011 Spanish General Elections." *Communication and Society* 27(1): 83–106. www.unav.es/fcom/comunicacionysociedad/en/www.comunicacionysociedad.com.

Appendix

Twitter Discourse Project Codebook

Overview

This coding manual is part of a broader project to evaluate presidential candidates' discourse on Twitter during their campaigns in Latin America and Europe. The main task you will be performing is coding different kinds of messages (i.e., frames) that candidates use.

Frames are a rhetorical device that speakers (in this case, politicians) use to communicate their ideas with a particular lens around them. For the purposes of this analysis, a frame is defined as a political actor imbedding meaning into a message by encouraging the listener to interpret an event or situation from a particular non-neutral perspective. Essentially, a frame is a way for politicians to convey information to their followers in a particular way.

This study looks to classify all the possible frames used by presidential candidates. While frames are the main topic of interest for this research project, there are also several other dimensions I'll be asking you to code.

Coding

Coding will take place in Redcap. The coding unit is a single Tweet. The Tweets will be "de-identified" to the extent possible; that is, ideally, you would not know which candidate sent the Tweet. In reality, however, that is not always possible—many Tweets mention specific cities or candidate names directly that will make it impossible not to know who is speaking (at the very least, what country they are from).

Furthermore, to code the Tweets, it may be necessary to view the media attached to the Tweet which will require you looking at the Tweet on twitter, thus exposing you to the speaker's identity. The primary concern is not that you know who is Tweeting, but that **even if you have previous knowledge of these candidates or countries more broadly, it is important to evaluate every single Tweet individually and without bias**. To help with this, the Tweets will be randomized between candidates and across dates.

The importance of looking at the media of the Tweet cannot be overstated: for example, one Tweet read:

"266: the number of jobs that Andres Manuel created as Head of Government."

From this alone, it is challenging if not impossible to understand if the speaker considers this a lot of jobs or not. However, if you navigate to the Tweet to see the image, you can clearly see that the speaker (in this case, Jose Antonio Meade of Mexico) considers 266 to be a very low number.

266: el número de empleos que creó Andres
Manuel como Jefe de Gobierno.

⟳ Translate Tweet

It also helps to look at the hashtags: those that refer to cities or locations could help you determine this is a campaign event (which helps you classify the function of the Tweet), while others may help you determine which frame to use.

What will be Coded (per Tweet)

Each Tweet will be coded according to 8 dimensions, each of which will be described below.

1) the master frame;

2) the relevant actors;

3) the sub-frame;

4) the difficulty in classifying the sub-frame;

5) the perceived strength of the frame;

6) the issue that the Tweet addresses; and

7) the function of the Tweet

8) whether the frame used positive, negative, or neutral language;

Some coding categories are dependent on your answers to previous coding categories. For example, the actor, master frame, sub-frame, difficulty classifying the sub-frames, and strength of the sub-frame are categories that are dependent on one another. For these categories, coding each Tweet in a particular sequence may make the identification of subsequent categories easier. However, sticking to a specific order is not absolutely critical. You may jump between categories as you decide on the proper categories. Other categories (the presence of positive or negative language, the issue, and the function of the Tweet) are independent: to code these categories, you

only need the Tweet itself because your answer to these categories is not dependent on your answer to other categories. Below is a brief description of each category.

How to Code the Tweets

Before you start coding the Tweets, it is critical to **read for subtext or, stated differently, to take a holistic view of the Tweet,** rather than taking the Tweet at face value. What is the candidate saying between the lines? <u>**Consider the Tweet as a whole before you start coding**</u>, before breaking it into the constituent parts required by the coding categories.

Tweets can only be 280 characters; such short texts mean the sub-frames may not be immediately clear, but by reading for subtext, taking the "spirit" of the Tweet as a whole into consideration, and *then* coding each category, you should have arrived at your conclusions systematically. Reading the text as a whole will specifically help you determine the master frame, the first coding category.

1) Master Frames

Master frames represent the highest level of aggregation. Conceptually, master frames refer to how people see the world in terms of who should hold power (Caramani 2017). There are 4 coding options for this category.

Master Frame	Who holds power (and who doesn't)?
1 = Type A: views the political world as a <u>divide between two groups</u>: "the people," who are understood to be virtuous and comprise a majority of the population, and "the elites," who are vilified for their self-interest and lack of representation of what the people want ("the will of the people"). **Explanation:** The antagonism comes from the idea that power legitimately stems from "the people," but "the elites" have taken this power and do not represent the "will" of "the people." Thus, "the people" and "the elites" are engaged in a struggle for power, and that struggle is inherently moral in nature. This type views the political world in terms of a virtuous group (the people) that has been wronged by the enemy group (the elites).	The people should hold power (over the elites)
2 = Type B: advocates for <u>power to be shared</u>: diverse interests are given voice, particularly from minority groups. **Explanation:** This emphasizes a more equitable form of power sharing. Some of the "key features and institutional structures that are intrinsic to [Type A]" include "compromise, mediating institutional bodies, and procedures that ensure, most notably, minority rights" (Akkerman et al. 2014, 1327). This type "sees political conflicts as struggles against impersonal forces rather than against diabolical groups and individual" (Busby et al. forthcoming, 2), in contrast to Type B.	Power is shared; at the very least, diverse interests are given voice (note: power does not have to be shared *equally*)
3 = Type C: prioritizes the <u>power of expertise</u>. Emphasizes practical applications and outcomes rather than ideals and focuses on the ability to deliver these outcomes. **Explanation:** Type C emphasizes practical applications and outcomes rather than ideals and focuses on the ability to deliver outcomes (generally because of experience) held by those delivering the	Those that can deliver favorable outcomes (specifically refers to politicians)

outcomes. The discourse does not frame issues in moral terms or paint them in black-and-white. Instead, there is a strong tendency to focus on narrow, particular issues. The discourse will emphasize or at least not eliminate the possibility of natural, justifiable differences of opinion. The discourse avoids a conspiratorial (moral) tone and does not single out any evil ruling minority.	
4 = Neutral: Master frames are those that do not fall into any of the above categories. Often, they cannot be classified into another master frame because they are missing a critical component of these other frames. **Explanation:** this is a "catch-all" category for frames that cannot in and of themselves be classified into just one of the above master frames. This is usually due to ambiguity—neutral frames can apply to a number of different world views, and this ambiguity necessitates its own category. Note that a neutral master frame does not imply that there is no bias, judgment, or moral component. These components are just not enough in and of themselves to indicate a master frame that fits into Types A-C above.	

2) the relevant actors;

This category refers to *who* the Tweet references (implicitly or explicitly): who is the one that is doing the action? Who is the one receiving it? Determining the actor will help to determine the precise frame. These are broad categories, and as a result, multiple interpretations exist. Which interpretation to choose may be in part derived from the master frame.

Once you've determined the master frame, identifying the actors will help you to determine the sub-frame.

Actor	Interpretation 1	Interpretation 2
"the people"	**1** = the good is embodied in the will of the majority, which is seen as a unified whole, perhaps but not necessarily expressed in references to the "voluntad del pueblo"; however, the speaker ascribes a kind of unchanging essentialism to that will, rather than letting it be whatever 50 percent of the people want at any particular moment. Thus, this good majority is romanticized, with some notion of the common man (urban or rural) seen as the embodiment of the national ideal. When using this classification, note that it should refer to the entire body of the candidate's supporters: all those he considers to be "the people." If only a sub-set (such as, for example, teachers, students, members of a particular town mentioned by name, then the appropriate classification is other, and to specify which group the speaker is referring to) *Explicit signifiers:* "el pueblo," "la gente," "Americans [or other nationality]," "the people" – this could also refer to specific townspeople ("the people of Merida") at campaign stops along the way	**2** = Democracy is simply the calculation of votes. This should be respected and is seen as the foundation of legitimate government, but it is not meant to be an exercise in arriving at a preexisting, knowable "will." The majority shifts and changes across issues. The common man is not romanticized, and the notion of citizenship is broad and legalistic. *explicit signifiers:* "citizens," "Mexicans [or other nationality]," "the people" *Implicit signifiers:* "we," "us" *Example:* "In this campaign, we are committed to listening twice as much as we talk. That is why my government will be the true government of the people [los ciudadanos in Spainsh], in which the needs of the citizens will be resolved. While this might seem like interpretation 1, it refers to citizens in terms of their will

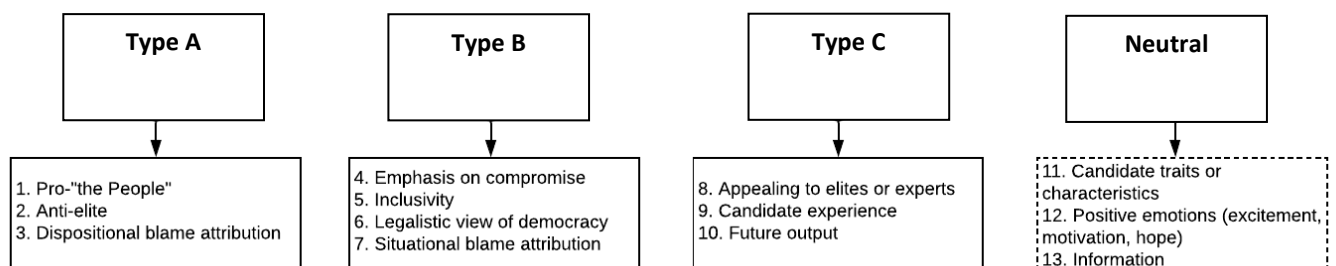|		
	Implicit signifiers: "we," "us"	
"the elite" (This most often applies to political elites)	**3** = The evil is embodied in a minority whose specific identity will vary according to context. Crucially, the evil minority is or was recently in charge and subverted the system to its own interests, against those of the good majority or the people. Depending on the context, political elites who are part of "the establishment" are often the primary target of politicians. *Explicit signifiers:* "the establishment," "the politicians," specific names of parties, other candidates, or individuals *Implicit signifiers:* "them," "they,"	**4** = The discourse avoids a conspiratorial tone and does not single out any evil ruling minority. It avoids labeling opponents as evil and may not even mention them in an effort to maintain a positive tone and keep passions low. Calling out an opponent for their poor performance could fall under this category—calling them evil or implying they intentionally harmed people would fall under Interpretation 1. *Explicit signifiers*: referring to other parties, or "incumbents" *Implicit signifiers:* "they," "them"
Other (usually, an in-group or out-group)	**6** = Generally refers to a specific [out] group (such as immigrants, or a particular ethnicity or race), but it does not necessarily explicitly identify this group (it may just be implied). The out-group does not necessarily have to be citizens of the country; it could be foreign entities (such as the United States). The important distinction is not the explicit identification of a group, but the implication that this group does not belong to "the people." *Explicit signifiers:* Referencing a specific group identity *Implicit signifiers:* "they," "them"	**7** = This may include reference to specific groups, generally in a positive sense of inclusivity and diversity. Since there is no romanticized notion of "the people," there is usually no out-group. In essence, an in-group refers to any subset of the overall population of the people as described in actor category 2 (the people, interpretation B). *Explicit signifiers*: Referencing a specific group identity (indigenous people, for example), students, teachers, members of a specific town *Implicit signifiers*: "They," "them"
The candidate, their party, or members of their party	**8** = This can refer to the candidate themselves, their party or party coalition, or other members campaigning under their party/coalition for other positions (not the presidency) *Signifiers:* "I," "we," name of party or other party officials	
The opposition	**9** = This includes any and all opposition candidates and their parties, and prominent members of the opposition parties (such as party figureheads, like former presidents) *Signifiers: name of party or other party official, name of candidate, references to other candidates/parties*	
No actors	**10** = Some frames will not have actors	
The media	**11 =** The media. This could refer to specific media personalities or media channels, radio stations, etc.	

There may be more than one actor per Tweet: if so, determine which is the primary actor and which is the secondary actor. In doing so, consider who is the Tweet really about? Who is the actor doing the action (primary

actor), versus who is receiving the action (secondary actor)? This is most often true when the actors include the candidate and some group of constituents. Is the Tweet specifically about the candidate and what the candidate hopes to achieve, or is the Tweet directed at a specific population they hope to serve? For example:

"We are going to shield the border so that US weapons do not enter Mexico and do not kill our people."	This Tweet references both we (the candidate/their party), and "our people". The primary actor would be the candidate, while the secondary would be the people.
Between the fracking to extract oil and the fumigation with glyphosate that will be from Colombia's water? I proposed that water is a priority for human consumption and food production and therefore there will be neither fracking nor glyphosate.	The primary actor would be the candidate, and there is no secondary actor. **Note that it's not just about the order in which the actors appear, but which actor the Tweet is really revolving around**--this one is about the candidate's position, which he is juxtaposing against the opposition's position; but even had the Tweet mentioned the opposition first, if the focus was on the candidate's proposal, then the candidate is still the primary actor.

3) Sub-frames

Sub-frames are the different ways in which the master frames manifest in rhetoric. Because the above master frames represent overarching worldviews, often they appear in partial form. **The sub-frame should match with the master frame it is nested under**: if you select Type A as a master frame, the sub-frames available are 1-3; if you select Type B, the options are sub-frames 4-7; Type C, 8-10; and neutral, 11-13. If you have a mis-match between the master and the sub-frame, go back to step 1 and re-evaluate the Tweet as a whole and see if either the master frame or the sub-frame is incorrect. If you are still stuck, flag it and we will go over it as a group.

Type A	Type B	Type C	Neutral
1. Pro-"the People" 2. Anti-elite 3. Dispositional blame attribution	4. Emphasis on compromise 5. Inclusivity 6. Legalistic view of democracy 7. Situational blame attribution	8. Appealing to elites or experts 9. Candidate experience 10. Future output	11. Candidate traits or characteristics 12. Positive emotions (excitement, motivation, hope) 13. Information

Master Frame	Subframe	Explanation	Example (s)
Type A	1 = Pro-"the people"	when a politician talks "in the name of 'the people', referring primarily to its will" (Cranmer 2011) The idea that the candidate is the "true representative" of the people (Engesser et al 2017) Puts the people's problems "at the core of the political agenda" (Casero-Ripolles et al. 2017, 990). The people are often characterized as hard-working (Engesser et al 2017)	If only for the will of the people we could say 'this rice has already been cooked', but we must prepare ourselves to face any fraud attempt. That's why I ask you to help defend the vote and democracy. We continue to collect the feelings and wisdom of the people. We are going to win the first day of July and we are not going to fail the people. Power only makes sense, and becomes a virtue, when it is placed at the service of others
	2 = Anti-Elite	Attacking anything that is "business as usual" or "how things have always been done." This is a pure and general form of anti-elitism, where "a political actor criticizes elites, such as political adversaries, the state, or the media" (Cranmer 2011, 293). It does not necessarily call out a specific elite actor, but it may.	In Tapachula, on the border of Mexico with Guatemala, I reaffirmed the commitment to banish corruption and govern with austerity. **There will be no gasolinazos**. **The Reformation, as emblem of the conservative press, fifi, is not able to rectify when it defames**, as it did yesterday with the supposed payment of MORENA of 58 million. In their code of ethics, the truth does not matter, but the interests and ideology they represent. Better we are free.
	3 = Dispositional blame attribution	Blaming some specified group of people for a particular failure-- allows actors to place the onus on particular elites or groups of people (such as immigrants) for specific failures (real or perceived) and for knowingly exploiting the interests of the people. Implies that elites/others knowingly exploited the interests of the people (Hawkins and Rovira Kaltwasser 2018, 7). This frame identifies political actors with agency.	Those responsible for the Hidroituango disaster after the genocide of the town of Ituango are two: Álvaro Uribe Vélez and Luis Alfredo Ramos: the complete degradation of the traditional political class of Antioquiapic. The PSOE has given a secret order to the ministries not to execute 50% of the budget. They bring us the cuts through the back door. It is the same as Montoro did and it means recovering the austerity policy of the PP. That is not the Spain you want. IN THE COUNTRY OF THE RAPES COMMITTED BY IMMIGRANTS He refused the stalker's advances

Type B	4 = Emphasis on compromise or cooperation	Stresses the importance or benefits of working with other groups or coalitions in the *political* arena. This frame may include references to coalition building, for example, or other references to governing with multiple groups.	The future is for everyone! Today I celebrate that the @MovimientoMIRA party supports our country project. We continue forming a coalition that will motivate Colombians so that class hatred no longer exists and so that from the differences we can build a better country.
	5 = Inclusivity	The discourse will emphasize the importance of the inclusion of groups, particularly those that are marginalized or disadvantaged. Rather than emphasizing a power sharing arrangement (like the above frame), it may simply mean giving these people a voice in some (often general or vague) capacity or listening to a group of people. More broadly, discourse may emphasize unity.	**We have a historical duty and commitment to our indigenous communities; As President, I will protect the rights of this population and we will work together to overcome their social backwardness.**
	6 = Legalistic view of democracy	Viewing democracy as the majority of votes (this is in contrast to the Type A frame of a romanticized "will" of the people). Tweets in this sub-frame may emphasize the duty to represent what the majority of the country wants (i.e., what 51% of the country wants, rather than what the group "the people" per Interpretation A want), or representing the country (or some subgroup) as a whole by meeting their usually broad, undefined needs (wherein specific needs being met would indicate an output sub-frame). Tweets in this category may reference listening to the people, but not acting on their will (which would indicate subframe 1)	More and more citizens are joining this project of future and certainty, which will result in **free, reasoned and conscious votes**. From now on I thank you. We will win! During the next three months, every week I will visit a family in their home. This time I visited Ana Laura, who invited me to eat with her husband and children. I want to listen to them and know what they think, leave me your messages and comments to be able to know them.
	7= Situational blame attribution	Situational frames tend to blame corruption/failed representation on "systemic causes such as globalization or technological change, and it tends to criticize rather than demonize political opponents" (Busby et al forthcoming, 8)	Mafia security has been broken in Medellin. Security does not depend on the number of deaths of young people. The Orion operation has failed. The wild posters of Mexico are taken to Medellin. I propose to integrate the youths to the university, the knowledge, the art and the Power
Type C	8 = Appealing to elites or experts	Deference to the expertise or opinions of organizations or actors outside the candidate or their party who have particular expertise (for example, government agencies or NGOs). This could include endorsements by actors outside the political party (but the endorsement should be described in a non-moralistic way). Appealing to a select group of individuals based on some attribute that they have, such as intelligence, wealth, or experience, operating on the belief that these individuals deserve particular influence. Endorsements by specific elite groups could be considered this type.	Fourteen entities commented on our environmental proposal. They evaluated these five criteria: water, climate change, deforestation, land use planning and new development models. I am touched by the support of Peter Singer, world-class philosopher, environmentalist and animalist. Peace with nature, respect for the animal, the other for us, for what is different from us, is the basis for humanity to live on the planet.

	9 = Candidate experience	refers to the candidate's unique ability to perform the job (or the opposition's inability to do so) Appealing to prior or current performance or particular attributes of the candidate or their party. This could take the form of talking about specific policy achievements, their years of experience in a position, their particular expertise on a subject area, their credibility in general, etc. It can also include announcing a cabinet or other appointment. This is the positive usage of this frame. The negative usage would be calling out an opponent/party because they lack experience or more broadly, they lack credibility.	Faced with a complex and uncertain global environment, Mexico needs a President with **proven international experience**. With the United States there is no room for improvisation. Here my editorial published today in the Arizona Republic
	10 = Future output	the projected output of a candidate—what is the candidate going to deliver if elected? With few exceptions, this category refers to promised policy outcomes, though it can also refer to positive consequences of electing the candidate or negative consequences of electing the opposition This frame can be used positively (as in the case of appealing to particular issues the candidate supports) or negatively (where the candidate criticizes his/her opponents for a particular issue stance)	In order to have transparency in the use of public resources, we will create a digital platform that, using blockchain technology, allows us to follow its course. Citizens will know exactly what money is allocated to, what it is used for and where it ends.
Neutral	11 = Candidate traits or characteristics	focuses on attributes or reputational considerations. Tweets where candidates are portrayed as "honest" or "hardworking" (as examples) are incorporated into this category, as are tweets that describe specific actions taken during the campaign (things like "candidate X did action Y").	I'm the only candidate from the Northeast. I need to defend my people! They file in a complaint that they had made against me saying that my titles are false or I put falsehoods on my page of my life. My studies are what I said. 'As a good teacher, Fajardo is seen as convinced and patient, perhaps certain that changes take time but arrive, without haste, without manipulation, without buying consciences, a sowing that I hope the fertile electoral harvest he hopes for.'
	12 = Positive emotions (hope, excitement, motivation)	rhetoric that conveys hope or excitement, or general motivation for the election Emotions such as hope or excitement are not in themselves indicative of a particular worldview, *especially during the course of an election in which candidates hope to inspire positive emotions among their supporters/try to gain new supporters*. You will likely encounter many motivational frames that aim to drum up support for their candidacy, but to be considered a specific master frame, the emotions must be used with another frame.	We are 15 days from the end of the campaign and **the mood of the people is growing as if it would burst with happiness**. Never in Ticul or Chetumal had we held such emotional and large meetings during the week.
	13 = Presentation of facts	This discourse is purely factual: it presents information, but does not impose a particular frame.	Another intense day of campaign: We talked with members of the Mexican Business Council; we present the

			environmental agenda in Zacatelco with Josefa González Blanco; We were in Apizaco and in Xalapa, Veracruz, accompanying Cuitláhuac García, our candidate for governor. I invite you to follow our press conference…

If you selected subframe 10 when you are coding, this is a follow up question that will be asked. Because this sub-frame can take many forms, please select one of the following options that best describes the output the speaker is talking about.

1 = Mention of a broad group of issues or a proposal, but not a specific issue	Extremely vague (if a policy is identified, there is no information provided about it)	"Look at our issue policies"
2 = Identification of a specific issue but not necessarily the intended outcome	Not fully specified: Policy X is identified, but Outcome Y is not	"If elected, I will enact policy X" (ex: if elected, I will reduce taxes on the middle class)
3 = Expressing a desired outcome but not the specific steps/policy to get there	Not fully specified: Outcome Y identified, but the specific policy X is not (i.e., it is not clear what steps the speaker will take to achieve the outcome)	" I want to improve/enact outcome Y" (ex: I want to improve education, health, etc.)
4 = Identification of a specific issue *and* the intended outcome	Fully specific: Policy X and the subsequent Outcome Y are identified	" I will enact policy X to accomplish outcome Y" (ex: I will enact a country-wide minimum wage to reduce income inequality in the countryside)
5 = Not applicable	Subframe chosen in the above section is not 10, output	

This is a self-reported measure of how difficult it was to identify the frame you selected. There are 3 possible values for this category:

0 = easy	Little to no uncertainty: actors were clearly identified; only one sub-frame seemed to apply
1 = somewhat challenging	some uncertainty: There were multiple possible frames, but one frame or sub-frame stood out
2 = very challenging	high level of uncertainty: There were multiple possible frames, and no frame clearly stood out as the predominant one

How close does this frame come in representing the *master frame*? This coding category requires you to read the Tweet for subtext and focus on 1) whether the critical elements from each master frame are present, and 2) whether these elements are mixed in with elements from other master frames or not.

2 = Strong. Comes extremely close to the **ideal master frame**, expressing all or nearly	Example(s): According to the survey of 'Saba' we grew after the debate. **They could not cheat us and that's why the dirty war**

all of the elements of the master frame, and has nearly elements that could be considered to represent a different master frame (if these other elements are present.	**intensifies**. Everything will be useless, nothing and **no one can stop the longing of millions of Mexicans for a change**. (Type C master frame, contains references to both the people and the elites) **The future is for everyone**! Today I celebrate that the @MovimientoMIRA party supports our country project. We continue forming a coalition that will **motivate Colombians so that class hatred no longer exists and so that from the differences we can build a better country** (Type B master frame, references both power sharing and the moral element of class hatred)
1 = Moderate. A Tweet in this category is **moderately reflects the master frame by including** some but perhaps not all identifiable elements of this master frame, and either does not use these elements consistently or tempers them by including elements from other master frames.	Example: **We continue to collect the feelings and wisdom of the people**. In the morning we were in Tequila, Jalisco, and in the afternoon in Compostela, Nayarit (Type C master frame, but also has informational elements, and doesn't reference the elites) My agenda is social, cultural and environmental. I am committed to the protection of the swamps and the páramos. I want all Colombians to protect the environment (Type D master frame, talks about a particular issue but it is vague in terms of referring to particular outputs or ways to achieve this)
0 = Neutral. A Tweet in this category is considered **neutral**: it uses few if any elements tied to specific master frames, or they cancel each other out. (Note: if you coded the master frame neutral or informational, this category should also be 0)	Example: Sunday full of joy in eastern Antioquia. On the street with young people who have already lived how #LaFuerzaDeLaEsperanza can transform society. We know that #SePuede govern with decency. See you in Marinilla, El Carmen and San Antonio de Pereira (neutral master frame, could be used with any master frame—nothing in it to indicate how power would be shared)

6) The issue that the Tweet addresses;[16]

What is the main topic of the Tweet? These categories are meant to be broad, but there are categories for "no issue" or "other" just in case a Tweet mentions something that does not fit easily into one of the following descriptions.

Subjects	Description	Example
1 = Economy	Tweets including subjects such as jobs, unemployment, salaries, deficit, public spending, debt, crisis, taxes, entrepreneurship, contracts, self-employed people, agricultural policy, and so on. This is a somewhat narrow category that should refer explicitly to the economic realm.	"+ 1 million jobs # since February 2014, of which + 53% on permanent contracts. Highest employment rate since the #Istat time series exists. # Youth unemployment at the lowest levels of the last 5 years."

---

[16] This category is adapted from Casero-Ripollés, Sintes-Olivella, and Franch (2017), but adapted for a smaller number of categories.

2 = Social policy	Tweets including subjects such as pensions, health, education, the welfare state, poverty, social justice, equality/inequality (including gender-based violence), housing, immigration, childbirth, drug rehabilitation, and so on. This is a broader category that encompasses some economic-adjacent issues (inequality, welfare) that affect people.	"To those under a certain income threshold, it could be an increase of 1000 euros a month for each dependent child, the State pays the necessary sum to arrive at a dignified life. The sum may vary depending on the area of the country where you live."
3 = Culture, media, and sport	Tweets including subjects related to cultural industries (cinema, literature, art, mainstream media, social media, etc.) and sport.	"The State must support our athletes! - The recognition of athletes like Carolina Marin, Saul Craviotto or Lydia Valentin cannot be a miracle. It must be guaranteed!"
4 = Science, technology, the environment, and infrastructure	Tweets including subjects related to research and development, network infrastructure (such as fiber optic, ADSL, or Wi-Fi), transportation infrastructure (railway, airports, roads, etc.), pollution, flora and fauna protection, climate change, and so forth.	"The planned future: the environment above all National event for the presentation of the # Environmental Program of the 5 Star MoVement."
5 = Terrorism, crime, and insecurity	Tweets related to terrorism in all its forms and crime/criminal activity or general concerns about insecurity.	"I will work hand in hand with the mayor of #Cali so that we can stop the exponential growth of many crimes in the city."
6 = Foreign affairs	Tweets alluding to the European Union, the United States, international relations, or other parts of the world.	"The United States also needs #Mexico. In my government, we are going to put all the negotiation issues on the table, and we will defend our country firmly on all fronts."
7 = Corruption and democratic regeneration	Tweets including subjects concerning political corruption and/or democratic aspects that need to be renewed or removed, like changes in electoral law, putting an end to the establishment and the privileges of the political class, and so on.	"The PSOE has given a secret order to the ministries not to execute 50% of the budget. They bring us the cuts through the back door. It is the same as Montoro did and it means recovering the austerity policy of the PP. That is not the Spain you want."
8 = Political strategy in office	Tweets including subjects concerning the intention of the **candidate if they were to win office** (i.e., not specific to the campaign period itself). For example, forming a certain type of government or possible (or impossible) government pacts/coalitions in the future. Additionally, if the candidate Tweets about	"Do you want to know all our government plan and know why so many people think that it is the most realistic, complete and successful proposal for Colombia? Here they find it

	multiple issue positions (the economy and social positions), classify it as political strategy.	complete. Read it and tell us what you think"
9 = Campaign organization and strategy	Tweets including subjects concerning the candidate **during the campaign period**. This can include questionnaires, surveys, information, analysis, and assessment of electoral results, or Tweets referring to the action of voting. It can also refer to Tweets about the running of the campaign and the organization of events, like rallies, meetings, political events, and media appearances by the candidates (more specific), or Tweets exalting the importance of party unity and exhorting sympathizers to join the party and earn victory (more broadly).	"In a week we will have an appointment with democracy. We will consolidate an arduous work that has taken me to travel the whole country, transmit my proposals and contrast capacity, preparation, honesty and responsibility with the other projects. With your vote, we will win" "Follow the first debate of candidates for the Presidency of the Republic."
10 = Immigration	Tweets about the topic of immigration	"Salvini at Tgcom24: 'Islam is a danger, stop at every presence'" "Elections 2018, Salvini defends Fontana on the immigration issue"
11 = Regional politics	Tweets relating to political subdivisions such as particular regions, states, etc. Note: this should not be used whenever a candidate talks about a particular city; it is more about the distribution of power within a country, such as the secession movement in Catalonia, Spain, or urban vs. rural politics.	"Mr. Sanchez, in Catalonia there are already enough competitions; what we need is that the people who manage them do so with loyalty to the Constitution." "In our program we propose formulas to improve the model of territorial organization. We want all Spaniards and Spaniards to enjoy the same rights, wherever they live."
12 = No subject or Other	Tweets that do not have a defined subject or that include expressions of courtesy (acknowledgments, etc.) or Tweets referring to the personal life of political agents. Tweets that cannot be placed in the above categories.	"I share this song, 'Cuidame tu', by Teresita Fernandez, played by Beatriz." "Happy Children's Day!"

What is the candidate trying to accomplish? Like the issue category, the possible functions are generalizable categories, with a residual category if needed.

Function	Description	Example
1 = Agenda and organization of political actions (including media appearances)	Tweets containing information on specific campaign actions in which the time and place are specified. **This should take place either in the near future, or be in progress at the time the Tweet is sent**. Tweets sharing links to a journalistic interview or TV show.	"This afternoon there is debate in the SBT. Do not miss it!" "Follow the first debate of candidates for the Presidency of the Republic." "Today at 7:00 pm there is an interview with Cyrus live on @recordtvoficial. Watch it!"
2 = Electoral program	Tweets on future political proposals or program proposals. This should be somewhat specific—not just vague intonations of making the country better.	"We have to increase competitiveness throughout the country. I propose to lower the VAT at the border and implement a National Infrastructure Plan to achieve prosperity in all states." "One of the key points of our program is less taxes for families and businesses. We will succeed in the Flat Tax, a single rate for all of 23% which will guarantee real economic growth, new jobs and a revival of investments."
3 = Management of political achievements	Tweets extolling or praising the achievements of the party and/ or leader. This could also include things like endorsements or responses to polls/early election predictions.	"Congratulations @diegosinhue! In #DebateGuanajuatense you showed that with responsible proposals, in this state we will continue to make good governments for the people. We will win!" "Thanks to Podemos, jobs are created and energy is saved, taking care of the planet."

---

[17] This category is adapted from Casero-Ripollés, Sintes-Olivella, and Franch (2017), but adapted for a smaller number of categories.

4 = Criticizing opponents	Tweets containing direct or indirect attacks on other candidates, political parties, other leaders (past or present) or other ideologies more broadly.	"Lopez Obrador is not change, it's just the opposite. Directly giving contracts to your friends is called corruption." "He was supposed to think about the Italians, but he thought only of himself. #Berlusconi spent 3,339 days in the government of the country and focused exclusively on his own affairs"
5 = Participation and mobilization	Tweets aimed directly at increasing support/votes **during the campaign**. This can include the mention of general campaign events (we were in XX city this morning), but the reference should be somewhat vague. Followers would not know where to go or what type of event based just on this Tweet alone (in contrast to function 1). Specific manifestation: requesting financial donations, encouraging people to vote for the candidate/party, or mobilizing volunteers. General manifestation: Tweets that contain inspirational messages about the campaign, or Tweets reinforcing the party values and containing concepts that identify the party, its ideology, or its values.	"<3 Vote for a big censure of corruption, inequality and political confrontation. Let's say it loud, very loud, voting for the Socialist Party. We are very close." "The second round opens up a golden opportunity: to win this election, an eye on the debate." "We are 15 days from the end of the campaign and the mood of the people is growing as if it would burst with happiness. Never in Ticul or Chetumal had we held such emotional and large meetings during the week."
6 = Personal life/ backstage or Manners/Protocol	Tweets where particularly the leaders show or talk about things from their private lives (leisure, hobbies, sport, etc.) or from backstage at political events or from the campaign. Tweets of thanks, sympathy, greetings, special occasions, and so on.	"Anyway at home, near my family in the warmth of our home! No better feeling! Thank you all for the expressions of affection that I could see on the way back and all over Brazil! A big hug to everyone!" "We continue with concern the fire in the cathedral of Notre Dame, in Paris. Let us hope that there will be no victims and that the firefighters will suffocate the fire, preserving this enormous jewel of heritage"

7 = Entertainment or Humor	Tweets encouraging community building around the party or the leader with an entertainment-based focus, or Tweets containing memes, jokes, or other humorous resources.	"Nothing better than ending Sunday with a good movie ... Defeating the dark side machines, you can!"
8 = Others	Tweets that cannot be placed in the above categories	

Does the candidate use mostly positive, neutral, or negative language? When considering this, think of the overall tone of the message, as well as the particular words used.

1 = predominantly positive language	"It's amazing how people are responding. Never have so many citizens participated as now in favor of real change. Look at Manzanillo." "I want to tell the country that I am honored that Dr. @ MoralesViviane gives us her support. With @mluciaramirez we are proposing a project for all Colombians, based on legality, entrepreneurship and equity, where we all fit."
0 = neutral language,[18] or equally positive and negative	"Conference with the international press. We are talking about climate change, fossil progressivism, new progressivism, anti-drug policy, the Venezuelan situation and the Middle East, which will be the new foreign policy of Colombia." "We have to eliminate the unnecessary expenses of the State. As president I will face the evasion; I will encourage investment and the formal hiring of workers, and I will contribute to improve their salaries."
-1 = predominantly negative language	"The real alliance: a scam to the Italians It passes a final majority report in the banks Commission thanks to 6 parliamentarians of the center-right who, upon leaving, reduce the quorum. Here is an advance from the government of mess-makers for which Renzi and Berlusconi work" "# SanchezMentiroso has been demonstrating for nine months that he lies more than he talks. Inside video"

---

[18] The use of the word "neutral" here is different than how it was used for neutral master frame. Here, neutral means there is no strong bias in the language.

You'll be coding many Tweets, so this brief description should provide justification about any items that required a judgment call. Since we will review each Tweet for discrepancies, this will help us to make the final determination about which code is most appropriate.

Examples:

I coded this as a Type B master frame because it emphasized power sharing and inclusion of voice--2 strong indicators of this type. I also coded it as an issue-based subframe because it talks about the specific proposals of students.

This was a neutral tweet that simply encouraged voters to vote for the candidate by using positive emotions and a reified sense of history. While there seems to be a vague reference to Type B, it's ultimately not enough to classify it as a master frame other than neutral (it's only vaguely implied, whereas the neutral subframes are fairly strong).

I coded this tweet as Type C 'trust in experience' because the candidate was talking about the woman he chose for his VP and the personal qualities and accomplishments that make her qualified. I put the issue as campaign organization and the function as participation and mobilization because they are explaining a new, important member of the campaign and hoping support increases because of her.

I coded this tweet as Type A 'pro people' because the party was lauding young people for their support and implying that young people are being driven to the party because it represents their ideals (patriotism, roots, etc). I put the issue as campaign organization and the function as participation and mobilization because the party was showing the support they have already gotten from the youth and explaining why they have that support in an effort to attract even more supporters.

Troubleshooting

What if there are multiple (sub)frames?

It is possible that more than one frame will be present in a single Tweet. Most often, that is going to be some reference to the people and the elite. There is a designated frame for this category: sub-frame #4, the people versus the elites. However, it is possible that there will be multiple frames in a Tweet. If that is the case, select a primary frame *and* a secondary frame. If you are unsure which frame is primary and which is secondary, designate the primary frame based on which frame the candidate devotes more attention. If for example there are 2 sentences about anti-elite, and only 1 sentence or a passing comment about an out-group, select the proper sub-frame for the anti-elite sentiment as the primary frame.

What if there are multiple issues referenced?

At times, Tweets (especially longer ones) will contain references to more than one issue (such as the economy and the environment, for example). If that is the case, the chances are that there is a deeper meaning behind the issues—the Tweet may mention multiple issues for strategic reasons (i.e., the real "issue" is political strategy